LA-UR-19-29270
Approved for public release; distribution is unlimited.

Title: Recommendations for Improving ASC Data Management

Author(s): Hick, Jason Cody

Intended for: Report

Issued: 2019-09-16

LOS ALAMOS NATIONAL LABORATORY
ADVANCED SIMULATION & COMPUTING PROGRAM
Recommendations for Improving ASC Data Management

Abstract
The Advanced Simulation & Computing (ASC) Program started in 1995 and has an essential mission that requires data stewardship for massive amounts of high performance computing data. This report provides general user requirements as determined through interviews of key users, describes the design attributes for high performance computing storage systems, and identifies improvements that would enable users to better manage their data.

Jason Hick
jhick@lanl.gov

Improving LANL ASC Data Management

Executive Summary

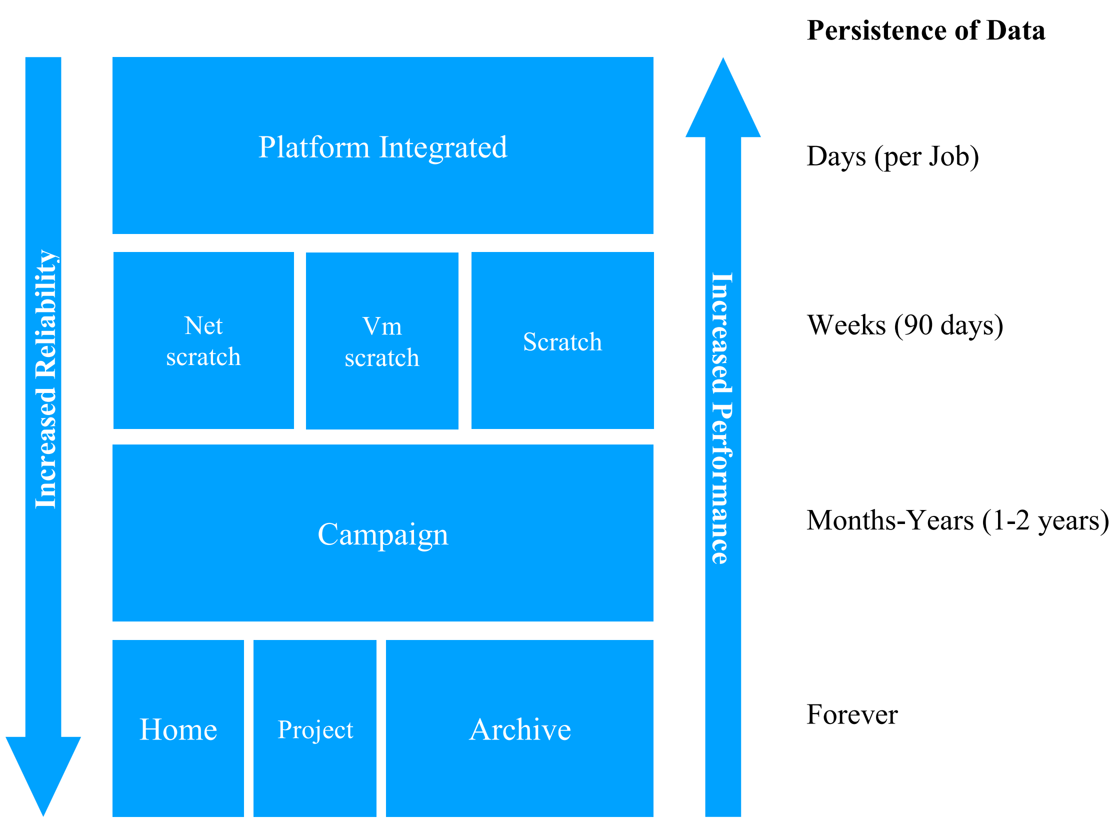
For the Advanced Simulation and Computing (ASC) Program at LANL, scientist productivity is a primary measure by which HPC storage systems are designed and managed. Storage market conditions and increasing storage requirements are forcing data management as a new measure by which future HPC storage systems will be designed and managed at LANL.

A series of user interviews revealed a number of specific storage use cases, requirements, and recommendations for improvement that are generalized in this report.

Historical usage shows that the top 10-20 users of a storage system at LANL own 90% of the data and are responsible for the data growth. For this reason, data management policy should focus predominantly on the needs of the top 10-20 users.

The final section of the report identifies recommendations for implementing changes to address gaps in enabling ASC user data management and shortcomings of existing storage systems. To maintain scientist productivity, it is important to ensure existing and future storage systems are capable of addressing the storage use cases.

Classes of Storage



Platform Integrated & Scratch Storage

Platform integrated storage is designed for providing the highest performance I/O to applications running on the largest systems at LANL. An example of platform integrated storage is DataWarp, deployed as a burst buffer for Trinity. This type of storage is expected to be deployed where data intensive workloads reside. In future systems, it is expected that platform storage will be high bandwidth and low latency and low capacity storage nearest the compute resources of HPC clusters. Reliability of platform integrated storage is centered around minimizing disruptions to running jobs, but generally lacking systemic protections. Data management of platform integrated storage is on an individual user basis and most often on a per-job basis.

Net scratch storage is designed for compute I/O where users are required to perform data management. There is no purge, rather users must manage allocations between themselves to ensure adequate application storage. Net scratch has general hardware protection against single device failures and avoiding single points-of-failure systemically, but avoids other layers of data protection to favor high bandwidth. Data management of net scratch is on an individual user basis.

Vm scratch storage is only available on the visualization cluster and is intended for animations and rendering using the Power Wall and CAVE. Vm scratch has general hardware protection against single device failures and avoiding single points-of-failure systemically, but avoids other layers of data protection to favor high bandwidth. Data management of Vm scratch is on an individual user basis.

Scratch storage is primarily designed for high bandwidth I/O to applications running on all systems at LANL. Performance degrades with high capacity utilization and thus purging is important to maintain performance. Purge is normally configured for 90 days. Scratch storage has general hardware protection against single device failures and avoiding single points-of-failure systemically, but avoids other layers of data protection to favor high bandwidth. Data management of scratch is on an individual user basis.

Campaign Storage

Campaign storage is for retaining data longer-term before down-selecting what important data to archive, or to bridge active data across computing campaigns. Performance is an order of magnitude better than archive and intended to enable moving entire data sets off scratch or project storage to alleviate capacity and data growth on higher performance storage (e.g. scratch). User data management will be required on this storage since there is no automated purge, but there is data expiration with each computing campaign.

Campaign storage has a high level of hardware protection, protecting against multiple simultaneous hardware failures such as multiple hard drives, failures of entire storage enclosures, and even protections across storage racks. With the data online, it is still susceptible to power events and as such users that desire the highest data protection are asked to ensure their most important data is archived.

Management of data within campaign storage is predominantly intended to be on an individual user basis, however, there are collaborative areas (e.g. visualization) that utilize campaign storage amongst a team of users.

Community Storage

Community storage is intended for enabling individuals or collections of individuals to store, organize and share their data, but is not designed to be a primary I/O resource. This is due to the technology choices favoring capacity and resilience rather than performance. Further, there is limited bandwidth available for any storage system in community storage.

Home storage provides for setting up and maintaining each user's custom environment as common across systems at LANL. It is designed with reliability as a key feature (nightly system backups, and enterprise-grade redundant hardware), but has poor performance capabilities. Users are asked to keep the minimum amount of data in home and restrict usage to serial I/O to prevent a single user from causing problems for other users. Management of data is intended to be on an individual user basis.

Project storage is intended for data that must be shared with team members for collaboration. It is capacity-oriented, not performance-oriented storage. Project storage is not purged. Users are required to manage their data in project storage to avoid ever-increasing capacity requests. Project storage has general hardware protection against single device failures and avoiding single points-of-failure systemically. Regular snapshots are provided to enhance protection against various events such as unintended deletion of data. Daily quota reports are provided to enable data management amongst a group of users.

Archive storage provides capacity-oriented storage for long-term data. Poor performance is relative to other classes of storage, but optimized for batch I/O rather than interactive I/O. The system works best retrieving larger amounts of data that it can sequence to minimize the number of removable devices it requires to fulfill the request. Though capacity-oriented, data is resident long-term. As such, users are asked to avoid a save-everything mentality, and only store what is important. The archive has general hardware protection against single device failures and avoids single points-of-failure systemically. In addition, the archive makes use of offline storage devices that provide protection against power incidents. Management of data is intended to be on an individual user basis.

Storage Use Cases

There are many ASC demands for HPC data storage at LANL. Rather than characterize every user or project, it is important to understand the general storage workflow of ASC users and identify any extreme demands on the data storage systems. Numerous meetings with ASC user representatives from the following areas characterized their use of HPC storage systems and identified a number of areas for improving their management of data.

Data Teams

Data Team representatives described the general workflow and needs for LANL nuclear data, atomic data, and equation-of-state (EOS) data. These teams focus usage on Project Storage and Archival Storage. Users generally have requirements for GB's of capacity with modest data growth (growing by a factor of 10 over 5 years), desire more reliable storage for long-term support, and have most performance requirements associated with heavy database usage (high numbers of small IOs) predominantly on read speeds (e.g. reading all data into memory for in-memory databases).

Attempts to utilize Platform Integrated Storage for high IO rates of the Data Teams were unsuccessful; IO rates were highest using Project Storage.

Code Teams

Representatives from the ASC code teams identified that their data requirements were in the 500GB to 6TB range per code in Project storage. There is significant growth in capacity requirements for code teams due largely to a move to new software for code repositories (e.g. GIT requiring copies and clones per user). Code team representatives identified that developer storage is a major issue in that there is no specific separate location, instead comingling developer storage within production code team storage. Much discussion centered around needing a more formal organization for Project storage for code teams around primary functions in order to improve data management:

- Software builds
- Software releases
- Regression and other testing
- Snapshots/backups
- Developer space

It became clear that having individual user quotas within team quotas would also be a useful feature for understanding when a particular user consumed too much capacity for the entire code team. There were several requests for running quota reports more frequently than once per day to enable code team leaders and users to see the effect of managing their data (e.g. not having to wait until tomorrow to see the benefit of removing unnecessary data).

Ownership of data is a significant issue amongst code teams, and the code team representatives expressed a desire to have ACLs in Project storage to enable need-to-know security access. Code team representatives lauded the current formal process of changing ownership of data for inactive users as working well for the code teams.

Special Use

Several methods to gather user requirements for Special Uses of storage were employed, to include meeting individually with big users, and gathering statistics on big users. In general, these are the most demanding users for all the storage deployed at LANL; they represent the 90% of the total data, and tend to saturate the performance of the individual storage systems.

Platform Integrated Storage and scratch are heavily used and very successful for enabling these special use computational jobs on the ATS-class systems at LANL. Special uses case users have requested Platform Integrated Storage be made available on CTS systems as well.

These are the largest single users of various scratch storage systems. Requests to extend purge intervals to reduce performance impacts of scanning scratch for purge candidates were requested. There were requests for enabling the equivalent of Project directories in the scratch storage systems to enable team collaborative storage.

Campaign storage is new to these users and only a few were aware of, or making use of campaign storage for their work.

Exercises

Exercises tend to utilize Scratch, Net scratch and Project storage from CTS-class systems depending on the exercise. Often the highest concern is successful execution of a time-sensitive job or series of jobs on a significant portion of the CTS-class system. In particular, exercises can launch thousands of jobs that wait on the same nuclear or EOS data. Storage resulting from these exercises may become forensic in nature and have specific long-term retention requirements. For satisfying retention requirements, users identified that they often use a combination of Project and Archival storage.

Contention for exercise storage and compute resources is limited by requesting a Dedicated Application Time (DAT), where HPC can reserve resources for the exercise.

When questioned about features desired, representatives expressed a desire for checksums for data transfers between storage systems to enable validation of data to be used for exercises. There was also a request for dual-copy on separate archival storage tapes to minimize data loss/corruption. There was a request to offer net scratch storage in the yellow network.

Exercise representatives mentioned that node locking failures are common when using Net scratch or Project storage for exercises.

Steps to Improve

It is important to enable use of keytabs in the yellow network, create a data movement system (e.g. yellow cap), and reduce the number of archives to 1 in the yellow/turquoise network. Today, these are impediments to better managing their data in the unclassified networks (yellow & turquoise).

In order to encourage use of platform integrated storage, HPC designed the HIO interface. It is intended to support use of DataWarp on Trinity by generalizing the protocol. An important step is getting HIO into FOUS production support in HPC.

HPC should explore the possibility of providing platform integrated storage for CTS-class systems due to the success of platform integrated storage for ATS-class systems.

Scratch will enable regular (weekly) purging to encourage good data management and improve overall bandwidth of the primary compute storage.

Project capacity is generally insufficient for supporting most use cases. HPC and users are coordinating to identify a limited capacity expansion through FY20. FOUS expects to plan a hardware refresh with capacity increase approximately every five years. Communicating this plan to users will help them manage their data to/from archival storage.

Project storage should consider offering a more formal structure for allocations around the functions of the code teams (build, release, regression/testing, developer) and tailor the snapshot configuration around supporting those different functions.

Campaign storage is prepared to grow significantly in the event usage increases to require more capacity. Currently, 30PB are deployed, with another 30PB in reserve. Bandwidth has been significantly improved since initially deployed with plans to increase further. HPC consultants need training on campaign storage to demonstrate its capabilities to users. We need advertising and encouragement of campaign storage at the management level (ADX and ASC) to increase utilization, as it was clear that not all users are aware of its existence or capabilities. Campaign requires a refresh plan approximately every five years to keep hardware viable, support costs reasonable, and capacity increases planned.

Archive storage currently lacks allocations or any limits on capacity. In order to discourage saving everything, the archive will implement allocations using quotas.

Next generation archival storage will explore improvements in performance by coupling more tightly with campaign storage.

New storage usage reporting will provide users a view of their data over time in different storage systems to assist users and managers in making decisions about data placement/movement and system growth planning. Reports need to be provided multiple times per day to improve user motivation for data management. A file indexing solution called Gufi is expected to improve data management by enabling some currently intractable mass metadata operations with existing storage systems, and in providing a more global view of where a user's data resides for improving their individual overall data management.

Propose peer review of top 10-20 users beyond the current ATCC proposal process. This would enable peers to periodically assess the value of retaining the data as a community. An effort similar to this achieved eliminating nearly 8 PB of the top user's data by ensuring derived data could be regenerated if required, thus deleting many old checkpoints.

Ensure scratch, project, and home storage supports code development or good interactivity of code developer data for compiling.

Gaps that Need to be Addressed

Many users have their own interfaces or techniques for moving data between the various storage systems. A common data transfer interface or file transfer solution between scratch, campaign, project, home, and archive systems would greatly increase user productivity in managing data.

Participation in the Remote Computing Enablement (RCE) tri-lab discussions is important to extend data management to tri-lab users.

Currently, LANL lacks a batch storage capability to offload large requests to move data, instead data transfers amongst storage systems are interactive.

There is no single tool that can provide a global view of a user's storage across systems at LANL. Instead, users are asked to manage their storage in a variety of distributed and independent storage systems.

Reports are not being provided frequently enough, currently once per 24 hours. In addition, data retention inside of snapshots needs to be addressed, such that a user does not have to wait 24 hours for snapshot release before freeing up storage. Currently, users tend to wait day(s) to reap the benefit of managing their data, thus associating data management with lost productivity.

Summary

This report identifies and recommends improvements for LANL ASC user data management. Interviews of users overall reflected a community of users that recognize the importance and vitality of the LANL HPC storage systems. Users were receptive to improving management of their data, but require tools and information that was not available to them in order to begin.

More specifically, the recommendations and gaps highlight the need for attention in the following areas:

- Identification, development, support of a common data transfer interface

- Changes to archival storage allocations
- Production tools for providing a global view of users data, reporting more frequently on user quotas, and simplifying platform integrated storage to applications
- Training and user education/awareness of campaign storage
- Increases in capacity for home/project storage
- Hardware refresh for plan for all storage systems

These improvements are contained nearly completely within LANL FOUS and will be addressed.